linguistic differences between China and North America or Western Europe where PROs tend to be developed. The objective of this study was to identify some of the linguistic and cultural differences between English and Mandarin as well as cultural differences between North America and Western Europe and China. **METHODS:** Five questionnaires relating to varied health states were reviewed and the cultural and linguistic issues arising during different stages of the translation process were examined. **RESULTS:** Numerous cultural and linguistic issues became apparent throughout the review, including: 1)Mandarin does not employ superlatives so "the worst" was translated as "extremely bad"; 2)some patients did not understand how to complete a VAS although the words themselves were clear; 3) there is no specific benefit of "disability compensation" in China—this is a general benefit which can include pensions; 4) Private clinics where respondents receive acupuncture and massage are very common in China so additional categories had to be added to a resource utilisation measure; 5) in Mandarin questions cannot begin with "how often;" instead they are phrased "does it often"; 6) with response options such as "never", "sometimes", appropriate responses can still be chosen; and 6) Low levels of obesity in China made recruitment for the linguistic validation of an obesity measure difficult. **CONCLUSIONS:** An increasing number of clinical trials take place in China. The issues raised above show some of the linguistic differences between English and Mandarin, and cultural differences between Western Europe / North America and China. These and other issues are important to consider when selecting, developing and translating measures for use in China.

PMC60

WILLINGNESS TO PAY FOR HEALTH INTERVENTIONS: BASED ON SYSTEMATIC REVIEW OF LITERATURES

Lee J[1], Park D[2], Ko S[1]
[1]Pfizer Korea, Seoul, South Korea, [2]Seoul National University, Seoul, South Korea

OBJECTIVES: The purpose of this study was to explore the incremental cost-effectiveness ratio (ICER) ranges of variant health interventions. **METHODS:** We performed the systematic review of cost-effectiveness studies about health interventions. Comprehensive literature review was conducted using MEDLINE and EMBASE with a publication date before March 24, 2009. We selected literatures based on the following criteria; 1) written in the English and Korean languages; 2) contained information on interventions relevant to diseases or health problems; 3) reported cost per year of life saved or contained sufficient information to calculate this ratio; 4) study methodology related to economic evaluations such as cost-utility analysis (CUA) and cost-effectiveness analysis (CEA), and 5) contained information whether the health intervention is cost-effective or not. We calculated the mean of reported ICER of major therapeutic area and major countries from all selected literature and extracted implicit ICER threshold based on the author's conclusion about cost-effectiveness. The value of ICER in each literature was converted to 2008 value reflecting exchange rate and price index. **RESULTS:** A total of 3662 articles were originally identified from the planned searching strategy. 1466 among them were finally included to calculate ICER range after abstract review and/or full text review. Two independent reviewers worked to select relevant articles and extract data. Mean incremental cost per Quality-adjusted-life-years-gained or Life-years-gained were US$130,514, US$25,446, US$30,448, US$70,420 for cancer, cardiovascular diseases, endocrine disease, and musculoskeletal disease, respectively. The ICER according to countries were also varied a lot. If we see the author's judgment on cost-effectiveness, some interventions were recommended as cost-effective even in the case ICER were approximately US$80,000, US$75,000, US$200,000 for UK, Canada, and USA. **CONCLUSIONS:** The judgment on cost-effective interventions could be different according to countries and disease. It tends to have higher ICER in severe disease and in developed countries.

PMC61

CATALOGUE OF EQ-5D SCORES FOR THE UK

Sullivan PW[1], Sculpher MJ[2], Ghushchyan VH[3], Slejko JF[3]
[1]Regis University, Denver, CO, USA, [2]University of York, York, UK, [3]University of Colorado Denver, Aurora, CO, USA

OBJECTIVES: The National Institute for Health and Clinical Excellence (NICE) has expressed a preference for the EQ-5D in cost-effectiveness analyses. Directly eliciting EQ-5D scores for specific disease states can be cumbersome and expensive. Use of disparate sources of utility scores by different researchers introduces variability and potential bias. Developing an "off-the-shelf" catalogue provides easily accessible EQ-5D scores from a single-source general population. The current study aims to develop a catalogue of EQ-5D scores for chronic conditions using U.K. preferences. **METHODS:** The Medical Expenditure Panel Survey (MEPS), a general population survey in the U.S., was pooled (2000, 2001, 2002 and 2003) to create a sample of 79,524 adults with valid EQ-5D responses. The EQ-5D was administered via a paper and pencil self-administered questionnaire (SAQ) to adults ≥ 18 years. UK tariffs for the EQ-5D (Dolan) were applied to questionnaire responses. Chronic conditions were classified by ICD-9 codes and Clinical Classification Category (CCC) codes. Ordinary least squares (OLS), Tobit, and censored least absolute deviations (CLAD) regression methods were used to estimate the marginal disutility of each condition controlling for age, comorbidity, gender, race, ethnicity, income and education. **RESULTS:** A catalogue of several hundred statistically significant marginal disutility (EQ-5D) scores for each chronic ICD-9 and CCC code are presented and compared. Marginal disutilities for each condition represent the isolated impact of the condition after controlling for comorbidity and covariates. **CONCLUSIONS:** This research provides a catalogue of "off-the-shelf" UK-based preference scores for the EQ-5D based on a single-source general population survey. Scores and marginal disutilities for a wide variety of chronic ICD-9 and CCC codes can be used to estimate QALYs in cost-effectiveness analyses in the UK. Although the preference function was derived from a U.K. population sample, there are limitations associated with using a US survey to develop the condition-specific questionnaire responses.

PMC62

ONLINE ADMINISTRATION OF THE TIME TRADE-OFF: A CASE-STUDY AMONG PEOPLE WITH RESTLESS LEGS SYNDROME

Aballéa S[1], Quilici S[2], Roïz J[3]
[1]i3 Innovus, Nanterre, France, [2]i3 Innovus, Uxbridge, Middlesex, UK, [3]i3 Innovus, Uxbridge, UK

OBJECTIVES: Time trade-off (TTO) utilities are traditionally elicited through face-to-face interviews, but these may be costly and time-consuming. We recruited a sample of restless-legs syndrome (RLS) sufferers using an online panel and administered an online version of TTO. **METHODS:** Members of a UK online marketing research panel were invited to participate, subject to confirmation of RLS diagnosis. The recruitment target was a minimum of 50 subjects in each of 4 severity categories based on the International Restless Legs Syndrome (IRLS) score. The TTO interface was adapted from a widely used, validated protocol for face-to-face administration. We explored associations between self-rated health according to IRLS, EQ-5D (UK tariff) and the online TTO. **RESULTS:** Over 15 days, 287 eligible participants (57, 90, 90 and 50 in mild, moderate, severe and very severe categories respectively) were recruited. Mean EQ-5D utilities by severity category were 0.80, 0.67, 0.51 and 0.29. 275 participants agreed to complete the TTO. Twenty-four percent were unwilling to sacrifice any life expectancy in their current state of health although they preferred a life in full health for a similar duration. Mean TTO utilities by severity category were 0.83, 0.82, 0.75 and 0.56. The correlation between EQ-5D and TTO utilities was 0.354 ($p < 0.0001$). TTO utilities were associated with EQ-5D utilities independently of IRLS score. Neither EQ-5D nor TTO utilities were significantly associated with age or gender, after controlling for IRLS score. Social class was significantly associated with EQ-5D, but not with TTO. **CONCLUSIONS:** Online recruitment enabled rapid recruitment of subjects with RLS. Correlations between TTO utilities and other variables were consistent with expectations, but the number of non-traders seemed high. Comparative studies with face-to-face TTO would be desirable, as the online approach seems an attractive solution to allow primary utility elicitation when time or cost constraints preclude a face-to-face survey.

PMC63

IS EXTENDING OF TIME TRADE-OFF EXPERIMENT TO 23 STATES PER RESPONDENT RATIONALIZED? ANALYSIS OF DATA FROM POLISH VALUATION STUDY

Golicki D[1], Jakubczyk M[1], Niewada M[2], Wrona W[1], Busschbach JJ[3], Macioch T[1], Hermanowski T[1]
[1]Department of Pharmacoeconomics, Medical University of Warsaw, Warsaw, Poland, [2]Department of Experimental and Clinical Pharmacology, Medical University of Warsaw, Warsaw, Poland, [3]Erasmus Medical Center, Rotterdam, The Netherlands

In first EQ-5D valuation studies based on time-trade-off (TTO) respondents valued 13 health states. **OBJECTIVES:** To evaluate the possible bias resulting from expanding the TTO experiment to 23 states per respondent. **METHODS:** In Polish EQ-5D valuation study, 7351 TTO valuations from 321 respondents were available before exclusions. First, we tested whether the mean valuation for each health state differed when it was valued in the middle of the experiment (6th-17th state) or at the end of the experiment (18th-23rd state). A series of t-tests for the equality of means and equality of variances was performed for each health state separately. To control for the multiple hypotheses testing the Hölm-Bonferroni correction was used. Moreover, to test the stability of regression coefficients within TTO experiment, Chow tests were performed: with the whole sample divided into two or three subgroups. In both cases the basic model with no interaction terms was used. **RESULTS:** The comparison of health state values when assigned during the middle of the experiment or at the end showed no statistically significant differences neither in mean nor in variance (the smallest P-values for means and variances comparison were equal to 0.0161 and 0.0060 respectively, with Hölm-Bonferroni threshold of 0.0023). Regression coefficients of two parsimonious models built on valuations 1st–17th ($n = 5009$) or 18th–23rd ($n = 1760$), did not differ significantly in Chow test ($p = 0.5521$). Similarly, regression coefficients of three parsimonious models built on valuations 1st–5th ($n = 1461$), 6th–17th ($n = 3548$) or 18th–23rd, did not differ ($p = 0.4334$), neither. **CONCLUSIONS:** As we found no systematic change neither in mean, nor in variance, nor in model parameters due to TTO experiment extension, we believe there is no risk of a bias or efficiency decrease in the model estimation. The present study supports the use of more health states per respondent in TTO valuations than previously thought.

PMC64

DIFFICULTIES WITH THE APPLICATION OF PRO MEASURES TO CULTURES OUTSIDE OF WESTERN EUROPE AND NORTH AMERICA IN MULTI-NATIONAL TRIALS

Furtado T, Houchin C, Wild D
Oxford Outcomes Ltd, Oxford, UK

OBJECTIVES: The majority of PRO measures are developed using a sequential rather than a cross-cultural approach and have been developed in the US and/or UK. When using these measures in multi-national trials, the assumption is made that the translated measure is acceptable for use in a multi-national trial, providing the translation